PRIVACY ON THE WEB:
Pulling Website Operators out of the Compliance Comfort Zone

Prof. Dr. Dominik Herrmann
Universität Bamberg

Slides: https://dhgo.to/dagstuhl19
Privacy: the claim of individuals to determine for themselves when, how, and to what extent information about them is communicated to others (Westin, 1970).

- **public information**
  - known to subject and public

- **private information**
  - known to subject only

- **partially private**
  - known to subject and some others

- **inversely private**
  - known to others, but not to subject

We lose control of our data.

http://m.cacm.acm.org/magazines/2016/7/204020-inverse-privacy/fulltext
The landscape of privacy solutions

unilateral protection

for nerds
e.g., browser extensions

bilateral protection

Why should I do more than absolutely necessary?
privacy as an issue of (legal) compliance

multilateral protection

for real nerds
e.g., Tor

- obtain information
- give consent
- proper processing
- subject access request
- account deletion
obtain information
give consent
proper processing
subject access request
account deletion
Before GDPR data protection officers had to provide a directory of procedures upon request to anyone. (§ 4g Abs. 2 BDSG a.F.)

„I hereby request the directory of data processing procedures according to § 4g Abs. 2 BDSG. Thank you in advance!“ (sent before GDPR)

<table>
<thead>
<tr>
<th>Type</th>
<th>Federal authorities</th>
<th>Universities</th>
<th>Businesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>210</td>
<td>425</td>
<td>200</td>
</tr>
<tr>
<td><strong>Reply received</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>request via e-mail:</td>
<td>49 %</td>
<td>51 %</td>
<td>48 %</td>
</tr>
<tr>
<td>request via letter:</td>
<td>47 %</td>
<td>56 %</td>
<td>57 %</td>
</tr>
<tr>
<td><strong>Directory provided</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>request via e-mail:</td>
<td>39 %</td>
<td>21 %</td>
<td>56 %</td>
</tr>
<tr>
<td>request via letter:</td>
<td>49 %</td>
<td>28 %</td>
<td>68 %</td>
</tr>
</tbody>
</table>
Anecdotal Results

 [...] According to BDSG § 4g par. 2, the data protection authority shall be provided with an overview of the information mentioned in § 4e sentence 1. Please **prove to me that you are the data protection authority.** *(misread the law)*

Due to the large number of data protection inquiries and the **significantly increased workload** due to the change-over to the GDPR, inquiries often cannot be answered without a **time delay.** *(after 4.5 months)*

[...] I am very pleased that now, at the end of the period of validity of the BDSG, **someone is actually interested** in our information according to § 4e sentence 1 no. 1-8 BDSG. [...]
obtain information

give consent

proper processing

subject access request

account deletion
How do providers react to subject data requests?
Evaluation of 120 websites and 150 apps

Informal request
Sehr geehrte Damen und Herren,
ich würde mich gerne darüber informieren, welche Daten Sie über mein Profil paul.meier@barmail.de gespeichert haben und wie diese verwendet werden. Bitte lassen Sie mir diese Informationen zeitnah zukommen.

Mit freundlichen Grüßen
Paul Meier

Formal Request
Sender: Paul Meier <paul.meier@barmail.de>
An: Data Protection Officer (or Customer Support)

Fake request
no acceptable answer within one week?

Sender: Paul Meier <paul.meier@fair-konsult.de>
An: Data Protection Officer (or Customer Support)

no acceptable answer within one week?
How do providers react to subject data requests?
Evaluation of 120 websites and 150 apps

Paul Meier                      paul.meier@barmail.de
Vogt-Kölln-Str. 30              Mobil: 0151-21512491
22527 Hamburg                   geb. am 05.03.1981

Sender: Paul Meier <paul.meier@fair-konsult.de>
An: Data Protection Officer (or Customer Support)

Hiermit fordere ich Sie gem. § 34 BDSG auf ... mir Auskunft zu erteilen, welche Daten über mich bzw. mein Profil paul.meier@barmail.de bei Ihnen gespeichert sind und zu welchem Zweck (§ 34 I-III BDSG i.V.m. § 6 II, § 28 Abs. 4) ...

Für die Erledigung setze ich Frist auf den [+ 7 Tage]. Nach fruchtlosem Ablauf der Frist behalte ich mir vor, die für Sie zuständige Aufsichtsbehörde gem. § 38 BDSG einzuschalten.
How do providers react to subject data requests?
Evaluation of 120 websites and 150 apps

Paul Meier
Vogt-Kölln-Str. 30
22527 Hamburg
Mobil: 0151-21512491
geb. am 05.03.1981

Paul Meier<paul.meier@barmail.de>

Informal request
Sehr geehrte Damen und Herren,
ich würde mich gerne darüber informieren, welche Daten
Sie über mein Profil paul.meier@barmail.de gespeichert
haben und wie diese verwendet werden. Bitte lassen Sie
mir diese Informationen zeitnah zukommen.

Mit freundlichen Grüßen
Paul Meier

Fake request
no acceptable
answer within
one week?

Sender: Paul Meier<paul.meier@fair-konsult.de>
An: Data Protection Officer (or Customer Support)

Sehr geehrte Damen und Herren,
ich würde mich gerne darüber informieren, welche Daten
Sie über mein Profil paul.meier@barmail.de gespeichert
haben und wie diese verwendet werden. Bitte lassen Sie
mir diese Informationen zeitnah zukommen.

Mit freundlichen Grüßen
Paul Meier
Results for subject access request (group: websites)

Fake request
- OK (0.18) 0.07
- no reaction (0.43) 0.02
- request denied (0.30)
  incomplete

Informal request
- OK (0.28)
- incomplete (0.28) 0.01
- no reaction (0.31) 0.07 0.05
  account vanished

Formal request
- OK (0.43)
- incompl.
- no reaction (0.23) 0.08 0.08
  not reachable denied

Reponses and amount of provided information vary a lot.

I've been dealing with your request, as this is only the second one we've processed it has taken a little while, please find the information requested attached.

Out of interest how often do you ask for these, and why?
Transparent?
obtain information
give consent
proper processing
subject access request
account deletion
complexity underestimated
Sehr geehrte Damen und Herren,

bitte löschen Sie meinen Account domi@exonet.de und bestätigen Sie mir die Löschung.

Grüße
Dominik Herrmann
Hallo,

hiermit bestätigen wir, dass wir Ihren Account gelöscht haben.

Es werden keine Daten an die Future TV Group weitergegeben.

Viele Grüße
wer-weiss-was Team

wer-weiss-was GmbH
Medienallee 19
85774 Unterföhring

w: http://www.wer-weiss-was.de
m: team@wer-weiss-was.de

Results from our experiment (2018):
In 35% of cases: account deletion was only possible manually.
Only 5% of those cases confirmed deletion by replying to the mail.
31.01.2019 08:38 - Dominik Herrmann schrieb:

Sehr geehrte Damen und Herren,

bitte löschen Sie meinen Account domi@exonet.de und bestätigen Sie mir die Löschung.

Grüße
Dominik Herrmann
obtain information
give consent
proper processing
subject access request
account deletion
The existing compliance regime has broken the principle of “giving consent”.

90% incomprehensible

86% too long

why?

This is often a lie ...

[...] provide any information [...] in a concise, transparent, intelligible and easily accessible form, using clear and plain language [...]
Giving Consent: Dark Patterns

European Data Protection Regulation

WE HAVE GOOD NEWS

Privacy is important for you ... and for us. A new regulation has come into effect in Europe regarding the way in which companies can use customer data. BQ intends to implement the said regulation and comply with it strictly. Consequently, we need your authorisation in order to stay in touch with you. :)”

- We'll send you personalised offers based on your profile, your geolocation or your devices usage data. No generic advertising!
- You'll only receive information of interest to you, such as exclusive promotions, gifts, new releases, trade-in schemes …
- We'll be able to offer you a better service, suited to your profile and needs.
- We won't forget about you: we'll tell you about anything that may interest you even if you no longer have a BQ device.
- You'll be able to revoke your consent at any time and unsubscribe without hassle.

Privacy is important for you ... and for us. A new regulation has come into effect in Europe regarding the way in which companies can use customer data. BQ intends to implement the said regulation and comply with it strictly. Consequently, we need your authorisation in order to stay in touch with you. :)”

- We'll send you personalised offers based on your profile, your geolocation or your devices usage data. No generic advertising!
- You'll only receive information of interest to you, such as exclusive promotions, gifts, new releases, trade-in schemes …
- We'll be able to offer you a better service, suited to your profile and needs.
- We won't forget about you: we'll tell you about anything that may interest you even if you no longer have a BQ device.
- You'll be able to revoke your consent at any time and unsubscribe without hassle.

See the privacy policy here

DECLINE

LATER ACCEPT

could have scrolled down!
Research Agenda: *How can we make it easier for users to exercise their rights?*

But: privacy is also **the right to be left alone.**

Users **may not want to be bothered** with it.

Happy to **delegate to proxies:** activists, the press, authorities.
obtain information
give consent
proper processing
subject access request
account deletion

matches reality? is it?
A sentence contained in many privacy policies:

Adresse mit anderen Daten von Google zusammengeführt. **Die IP-Adressen werden anonymisiert, so dass eine Zuordnung nicht möglich ist (IP-Masking).** Um die Nutzung unserer Webseite statistisch zu erfassen und zum Zwecke der

```html
<script>
(function(i,s,o,g,r,a,m){i['GoogleAnalyticsObject']=r;i[r]=i[r]||function(){
(i[r].q=i[r].q||[]).push(arguments)},i[r].l=1*new Date();m=s.createElement(o||'script');m.async=1;
m.src=g;m蛱iews=document.getElementsByTagName('script')[0];a.async=1;a.src=g;m蛱iews.parentNode.insertBefore(m,m蛱iews.nextSibling);
 ga('create', '********', '********'); ga('send', 'pageview');
</script>
</head>
</html>
```
Just an exceptionally seldom anomaly?

"Die IP-Adressen werden anonymisiert, so dass eine Zuordnung nicht möglich ist.

First 100 hits:
48 do not use Google Analytics
25 use it but without IP Masking

Ongoing study (1.3 million DE pages):
ca. 13% without correct IP Masking

Um die Nutzung unserer Webseite statistisch zu erfassen, schließen wir ein Tracking-Tool ein. Die Verwendung der Google Analytics soll jedoch ineffektiv (oder in falscher Reihenfolge) sein.

(function(i,s,o,g,r,a,m){i['GoogleAnalyticsObject']=r;i[r].q=i[r].q||[];i[r].l=1*new Date;i[r].q.push(arguments);},i[r].l=1*new Date;i[r].l=1*new Date;var m=s.getElementsByTagName(o)[0];a.async=1;a.src=g;m.parentNode.insertBefore(a,m);var ga=window.dataLayer=window.dataLayer||[];ga.push({'create': 'UA-XXXXX-Y', 'auto': true});ga('send', 'pageview');ga('set', 'anonymizeIp', true);</script>
So let’s just send over the 1.3 million URLs to the ca. 16 Data Protection Authorities in Germany? No.

DPAs have insufficient resources.

“www.hamburg.de has US-based trackers” sent to DPA on 12 July 2018.

First response in Jan 2019.

(June 2019: not fixed yet)

legally acceptable ≠ ethically acceptable
obtain information
give consent
proper processing
subject access request
account deletion

unilateral protection
bilateral protection
multilateral protection

How to break the asymmetry?
Publicity and Comparability

culture of outrage emotions

e.g. pride or enviousness

Naming and shaming to incentivize operators?
Existing site scanners focus on single sites, which does not provide sufficient context.
Compare Websites with PrivacyScore

PrivacyScore allows you to test websites and rank them according to their security and privacy features.

Create new site list

— or scan a single site immediately —

URL, e.g. privacyscore.org

Public benchmarks: incentive für site operators to improve their procedures

Crowd soucing: users can upload lists of sites and influence the ranking

Open Source (GPLv3+) and Open Data

Objective: repeatable studies with low effort
German Institutions of Higher Education
Submitted 6 months ago by foobar123

- sites passed all checks ................. 0
- sites with issues ...................... 331
- sites failed an area .................. 3
- sites with critical issues ............. 0
- Total number of sites ................ 426
- Sites with scan errors .............. 120
- No rating (missing data) ............. 3
- Excluded from scans ............... 0

HTTPS available: 79%
HTTPS used by default: 45%

Sites with leaks: 5%
Sites with known trackers: 30%

Average trackers per site: 56

<table>
<thead>
<tr>
<th>Rank</th>
<th>URL</th>
<th>Name</th>
<th># Trackers</th>
<th>Attacks</th>
<th>EncWeb</th>
<th>EncMail</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><a href="http://www.uni-regensburg.de">http://www.uni-regensburg.de</a></td>
<td>University of Regensburg</td>
<td>0</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
</tr>
<tr>
<td>2</td>
<td><a href="http://www.tum.de">http://www.tum.de</a></td>
<td>TU Munich</td>
<td>3</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
</tr>
<tr>
<td>3</td>
<td><a href="http://www.uni-northern.de">http://www.uni-northern.de</a></td>
<td>University of Northern Bavaria</td>
<td>14</td>
<td>🟢</td>
<td>🟢</td>
<td>🟢</td>
</tr>
</tbody>
</table>
Old data! You are looking at historic results. View latest scan for this site.
**Tracker**

**BY SITE OPERATOR AND THIRD PARTIES**

**Third Parties**

- **3 known trackers** (de.ioam.de, double...)
  - https://de.ioam.de/tiio?st=heise&cp=homepage&sv=ke&pt=CK&ps=lin&er=N2f=&r2=...
  - https://securepubads.g.doubleclick.net/gpt/pubads_impl_rendering_2019041601.js
  - https://cdn.mateti.net/mcp/onsite.min.js

- **No Google Analytics**

- **10 other third parties** (youtube.com, google.com...)

- **5 third-party cookies**

**Site Operator**

- **No first-party cookies**

- **Webserver in GDPR-implementing country: Germany**

- **Mailserver in GDPR-implementing country: Germany**

- **Webserver and mailserver are located in the very same country**

- **Privacy-preserving Referer header present**

---

**Mockup of next version**

---

**Old data! You are looking at historic results. View latest scan for this site.
Tests of PrivacyScore (some still in development)

Privacy and Tracking
- Third Parties
- Known Trackers
- Server Locations
- Cookie Syncing
- Browser Fingerprinting
- Subresource Integrity

Encryption Web Server
- HTTPS/STARTTLS available?
- Certificate: valid/ key length
- Insecure Protocols: SSLv3, ...
- Common Vulnerabilities: Heartbleed, ...
- HSTS
- Mixed Content automated Redirection to HTTPS

Encryption Mail Server

Miscellaneous Attacks
- Outdated Software
- Security Header
- Info Leaks

Tests of PrivacyScore (some still in development)
We have run PrivacyScore for 2 years.

almost 125,000 sites

> 1,700,000 scans

about 250 lists

banks, health insurance, schools, cities, political parties, politiciens, EU authorities, Alexa Top 500, data protection authorities, medical associations, chambers of commerce and industry, kids’ pages, retailers, booking portals, pharmacies, broadcasting agencies, news sites, universities, car makers, ...
Anecdotal evidence: some site operators are listening

<table>
<thead>
<tr>
<th></th>
<th>14 Aug</th>
<th>27 Oct</th>
<th>Delta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piraten</td>
<td>0</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Linke</td>
<td>0</td>
<td>1</td>
<td>!!</td>
</tr>
<tr>
<td>Die PARTEI</td>
<td>0</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>CDU</td>
<td>1</td>
<td>1</td>
<td>–</td>
</tr>
<tr>
<td>Grüne</td>
<td>1</td>
<td>2</td>
<td>!!</td>
</tr>
<tr>
<td>SPD</td>
<td>1</td>
<td>0</td>
<td>+</td>
</tr>
<tr>
<td>FDP</td>
<td>2</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>AFD</td>
<td>4</td>
<td>4</td>
<td>–</td>
</tr>
<tr>
<td>CSU</td>
<td>5</td>
<td>38</td>
<td>!!</td>
</tr>
</tbody>
</table>

Update 30 Jan. 2019
only one tracker on csu.de
On our homepage ... personal data is stored ... I am not amused [and] ask you to refrain from such tests in the future ..., otherwise I will [file] charges according to §202a of the criminal code.

Operation is legally acceptable in Germany.

arxiv.org/abs/1705.08889
GI INFORMATIK 2017
Well, then maybe you should really fix your site...

Addressing ethics of „dual use“

- No convenient search function for vulnerable sites
- Rate limiting to prevent Denial of Service attacks
- Blacklisting upon request (this is visible in the results)
- responsible disclosure when we find vulnerabilities (todo)

„Your site helps the attackers“

legally acceptable ≠ ethically acceptable
unilateral protection

bilateral protection

Research Agenda:
Transparency to break asymmetry?
Effective naming and notification?
How can we help authorities?

What is missing: Transparency enhancing technologies for authorities, press, and activists.
**PRIVACY ON THE WEB:**
Pulling Website Operators out of the Compliance Comfort Zone

---

**Fake request**

<table>
<thead>
<tr>
<th>Rank</th>
<th>URL</th>
<th>Name</th>
<th># Trackers</th>
<th>Attacks</th>
<th>EncWeb</th>
<th>EncEmail</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><a href="http://www.uni-regensburg.de">http://www.uni-regensburg.de</a></td>
<td>University of Regensburg</td>
<td>0</td>
<td><img src="icon.png" alt="Icon" /></td>
<td><img src="icon.png" alt="Icon" /></td>
<td><img src="icon.png" alt="Icon" /></td>
</tr>
<tr>
<td>2</td>
<td><a href="http://www.tum.de">http://www.tum.de</a></td>
<td>TU Munich</td>
<td>3</td>
<td><img src="icon.png" alt="Icon" /></td>
<td><img src="icon.png" alt="Icon" /></td>
<td><img src="icon.png" alt="Icon" /></td>
</tr>
<tr>
<td>3</td>
<td><a href="http://www.uni-northern.de">http://www.uni-northern.de</a></td>
<td>University of Northern Bavaria</td>
<td>14</td>
<td><img src="icon.png" alt="Icon" /></td>
<td><img src="icon.png" alt="Icon" /></td>
<td><img src="icon.png" alt="Icon" /></td>
</tr>
</tbody>
</table>